PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e20486
doi: 10.3897/tdwgproceedings.1.20486

OPEN ACCESS

Conference Abstract

Darwin Cloud: Mapping real-world data to Darwin Core

John Wieczorek[‡,§], Paul J. Morris[§], James Hanken[§], David B. Lowery[§], Bertram Ludäscher[|], James Macklin[¶], Timothy McPhillips[|], Robert A. Morris[#,§], Qian Zhang[|]

‡ Museum of Vertebrate Zoology, University of California, Berkeley, United States of America
§ Museum of Comparative Zoology, Harvard University, Cambridge, MA, United States of America
| University of Illinois Urbana-Champaign, Champaign, United States of America
¶ Agriculture and Agri-Food Canada, Ottawa, Canada
# University of Massachusetts, Boston, Boston, United States of America

Corresponding author: John Wieczorek (tuco@berkeley.edu), Paul J. Morris (mole@morris.net)

Received: 21 Aug 2017 | Published: 21 Aug 2017

Citation: Wieczorek J, Morris P, Hanken J, Lowery D, Ludäscher B, Macklin J, McPhillips T, Morris R, Zhang Q (2017) Darwin Cloud: Mapping real-world data to Darwin Core. Proceedings of TDWG 1: e20486. https://doi.org/10.3897/tdwgproceedings.1.20486

Abstract

Since its ratification as a TDWG standard in 2009, data publishers have had to struggle with the essential step of mapping fields in working databases to the terms in Darwin Core Wieczorek et al. 2012 in order to publish and share data using that standard. Doing so requires a good understanding of both the data set and Darwin Core. The accumulated knowledge about these mappings constitutes what we call the "Darwin Cloud." We will explore the nature of data mapping challenges and the potential for semi-automated solutions to them. Specifically, we will look at the "Darwinizer" actor and its usage in related workflows within the Kurator data quality framework and the implications for community-managed vocabularies.

Keywords

Biodiversity Informatics, Darwin Core, Data Quality

© Wieczorek J et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Presenting author

John Wieczorek

Funding program

NSF DBI 1356438 and 1356751

References

- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7 (1): e29715. https://doi.org/10.1371/journal.pone.0029715